ABSTRACT OF THE DISCLOSURE

A system encodes a separate assist channel that carries only a small amount of additional information in a hardcopy document to compensate for failure of an OCR system to accurately reconstruct a scanned electronic version of the hardcopy document. The assist channel is encoded by mapping data from a primary channel of the document into a plurality of groups. The plurality of groups is used to classify the symbols of the primary channel by the likelihood that a character will occur and the likelihood that the symbols of the primary channel will be confused during OCR processing.